Structure of a Traditional Baseline Data System

HE 1.1.1

December 1976

Sponsored jointly by

Alberta
ENVIRONMENT

Environment Canada
Environnement Canada

Penthouse, Jarvis Building
9925 - 107 Street
Edmonton, Alberta T5K 2H9
Phone 403/427-3943

91-03123/

ALBERTA OIL SANDS ENVIRONMENTAL RESEARCH PROGRAM

RESEARCH REPORTS

These research reports describe the results of investigations funded under the Alberta Oil Sands Environmental Research Program, which was established by agreement between the Governments of Alberta and Canada in February 1975. AOSERP is a 10-year Program designed to direct and coordinate research projects into the environmental effects of the development of the Athabasca oil sands in Alberta.

Enquiries pertaining to the Alberta-Canada Agreement or requests for copies of the research reports should be directed to:

Alberta Oil Sands Environmental Research Program
Penthouse, Jarvis Building
9925 - 107 Street, Edmonton, Alberta T5K 2H9
Phone (403) 427-3943.

Structure of a Traditional Baseline Data System

Sub-project HE 1.1.1

AOSERP Report 3

LETTER OF TRANSMITTAL

The Hon. D. J. Russell
Minister of the Environment
222 Legislative Building
Edmonton, Alberta

and

The Hon. R. LeBlanc
Minister of Fisheries and Environment
Parliament Buildings
Ottawa, Ontario.

Sirs:

Enclosed herein is the report on "Structure of a Traditional Baseline Data System" for the Alberta Oil Sands Environmental Research Program study area.

This report was prepared for the Human Environment Technical Research Committee of the Alberta Oil Sands Environmental Research Program, under the Alberta-Canada Agreement of 28 February 1975.

Respectfully,

W. Solodzuk, P. Eng.
Chairman, Steering Committee, AOSERP
Deputy Minister, Alberta Environment

J. S. Tener, Ph.D.
Member, Steering Committee, AOSERP
Assistant Deputy Minister
Environmental Management Service
Environment Canada.

Traditional Baseline Data System

DESCRIPTIVE SUMMARY

ABSTRACT

The research was to determine whether appropriate data exist for development of a comprehensive statistical baseline data system for the Athabasca Oil Sands region. The objectives included: a review and assessment of data, including the Statistics Canada Census; evaluation of weaknesses and deficiencies in existing data sources; recommendations based upon the foregoing evaluation; and a bibliography.

As an initial step, the existing data sources pertinent to the AOSERP target area were reviewed and assessed. The following criteria were selected to assist the evalution of data: the type collected, source, detail, geographic identification, accessibility, and time frame. These criteria permitted the assessment of whether the data would be amenable to geographically-coded, continuous monitoring systems.

It was found that the Statistics Canada Census provided the most detail, the most complete coverage of the target area, the smallest statistical areas, the greatest consistency in data and data collection, and the most regular collection. The local agency collection efforts were generally oriented toward specific goals and the data intended primarily for intra-agency use. In several instances, data were collected irregularly and were site-specific. The smallest statistical areas utilized were minor civil sub-divisions. While these statistical units are relatively comparable over time, they may be too large to be of value to a common small area system. Generally, data collecting agencies did not use coterminus boundaries; for example, health units, polling districts, school divisions, manpower regions, travel and industry zones were not comparable to one another. In many instances, these data collection areas were too large, extending beyond the entire AOSERP study region.

The recommendations of the report are:

1. Of all the data inputs reviewed, the Statistics Canada Census provided the most comprehensive coverage of the target area. Therefore, it is recommended that Statistics Canada Census data be given primary consideration in the initial development of the baseline data system.

2. Few of the statistical units reviewed were comparable to one another or historically comparable over the last census decade. It is therefore recommended that as a prerequisite to the establishment of a baseline data system, a common small area scheme be permanently delineated for the AOSERP target area.

3. It is further recommended that it would be premature to commence development of a baseline data system until the following conditions are satisfied:

    a. That a common small area scheme be adopted by local Provincial agencies and departments charged with the responsibility of data collection.

    b. That all demographic, social, economic, and biophysical data collected by local agencies and departments, as well as Statistics Canada, be geographically referenced to the common small area system.

It should be noted that while a baseline data system can be developed given the present status of data and statistical areas, its utilization would be hampered by the unevenness of data availability and lack of coterminus boundaries among data collecting agencies.

BACKGROUND

The purpose of a Traditional Baseline Data System was to facilitate the development of a comprehensive body of knowledge on human environment in the Athabasca Oil Sands region. The system was to be able to monitor various aspects of social change stimulated by the resource development. In the long run, the system could also facilitate an overall data integration by interrelating the socio-economic and biophysical research findings of the Technical Research Committees of AOSERP.

As a prerequisite to the development of a data system, it was necessary to determine the availability and the quality of currently existing socio-economic data pertinent to the region. Were the data available in a form suitable for routing into a system, the task of generating baseline data would have been greatly reduced. At the same time, effective relationships between the existing socio-economic and demographic data could have been established, thus laying foundations for subsequent research. If the data inputs were determined to be inadequate or unsuitable for establishing a data system, alternative methods for generating and integrating the baseline data would need to be developed.

Given the findings of the report on Traditional Baseline Data System, it is clear that because of the poor and uneven quality of existing

data, development of an operational data system is not yet possible. Moreover, because of the rapid in-migration of people and the fast pace of changes in the oil sands region, various kinds of data gaps have developed in the existing sources.

This situation has several implications for the research strategy of the Human Environment TRC. First, if the Committee wishes to use quantitative analysis, it will have to undertake its own systematic collection of primary data for the study region. Second, the data collecting agencies may need to be encouraged to collect the data in a more systematic way. Third, it would appear that one of the more promising research approaches would focus on predetermined areas of social concern; the data would be generated to answer problems specific to given research areas; and the incompatibility of data inputs across the research areas need not hamper the development of a comprehensive body of knowledge on the region. While a baseline data system could have been useful in facilitating this type of research, its establishment is not imperative to the rigorous examination of social change associated with the oil sands development. In the absence of a baseline data system, however, the overall integration of socio-economic and biophysical data in AOSERP may have to be limited to inter-disciplinary research projects.

ASSESSMENT

The Human Environment Technical Research Committee has reviewed and accepted the report on Traditional Baseline Data System, which was prepared by Dr. W. W. McVey of the Population Research Laboratory, The University of Alberta. The Committee has also accepted the recommendations contained in the report, but their implementation may not be feasible at this time.

The report is considered to be fairly comprehensive. It includes such documents as: survey response evaluation; census data file listings from 1961 to 1971; demographic, social, economic, and biophysical data contained in the inventory of statistical files in the Alberta Government; data available for Fort McMurray; a review of existing data bank systems; a bibliography of Syncrude Canada Ltd. reports; a list of respondents and questionnaire; and a bibliography of relevant literature.

In view of the unquestionable value of these documents, the Human Environment Technical Research Committee, in agreement with the Oil Sands Environmental Study Group, recommended that the report be made public and available to the AOSERP researchers as soon as possible. By identifying the existing data sources, the Report on Traditional Baseline Data System represents an excellent working document which ought to serve well as a basis for research on the human environment in the Athabasca Oil Sands region.

Paul W. Conway, Chairman
Human Environment Technical Research Committee

STRUCTURE OF A TRADITIONAL BASELINE DATA SYSTEM

by

Dr. WAYNE W. McVEY
Population Research Laboratory
The University of Alberta
Edmonton, Alberta

for

HUMAN ENVIRONMENT TECHNICAL RESEARCH COMMITTEE
ALBERTA OIL SANDS ENVIRONMENTAL RESEARCH PROGRAM

SUB-PROJECT HE 1.1.1

December 1976

alberta canada

ALBERTA OIL SANDS ENVIRONMENTAL RESEARCH PROGRAM
STUDY AREA

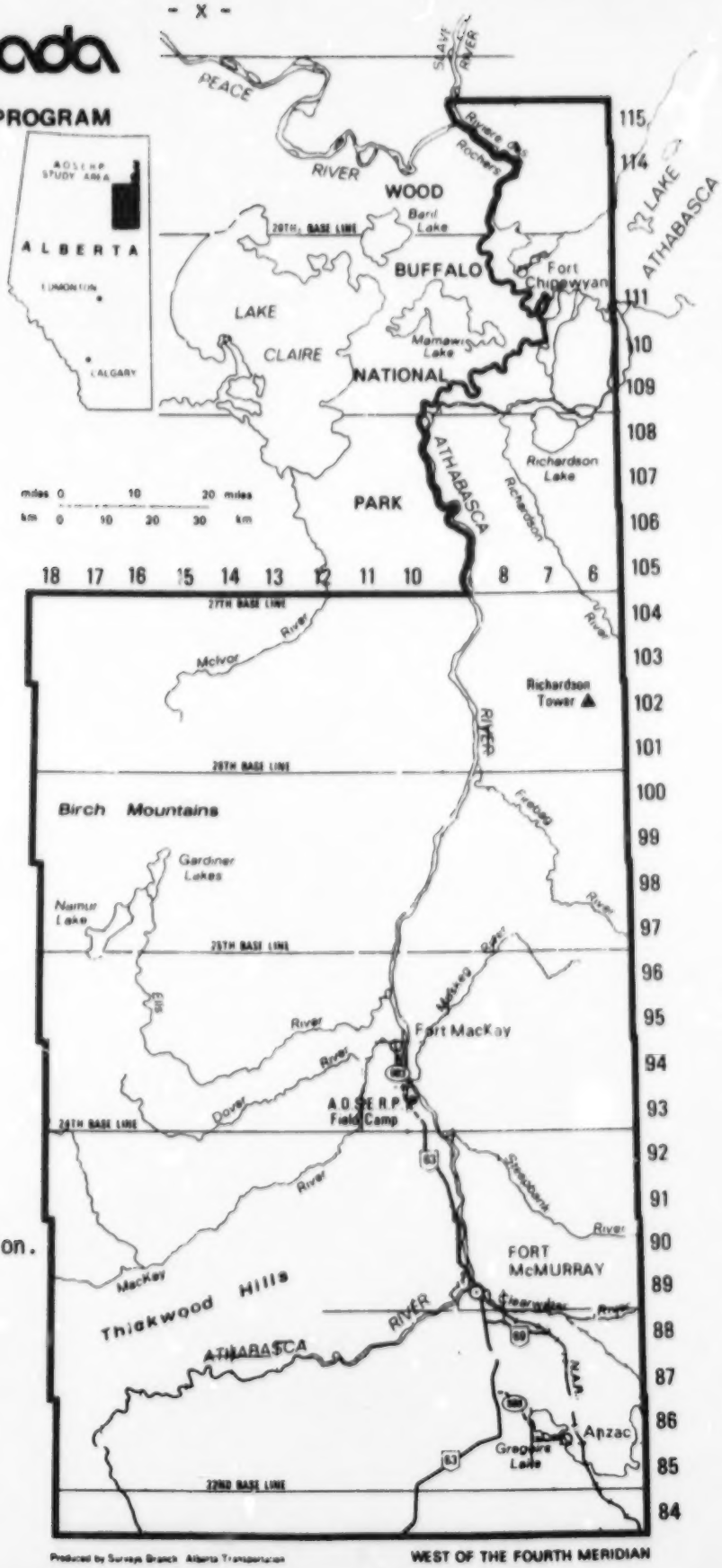

FIGURE 1.

AOSERP study area location.

CONCLUSIONS AND RECOMMENDATIONS

The following conclusions and recommendations are based upon an overall assessment of the present status of statistical areas and available data pertinent to the AOSERP target area. Of all the data sources available, it is evident that the Census provided the most detail, completeness in coverage of the target area, the smallest statistical areas, greatest consistency in data and data collection, and regularity in collection. It was found that local agency collection efforts were generally oriented toward specific goals and the data intended primarily for intra-agency use. In several instances, data were collected irregularly and were site-specific. The smallest statistical areas utilized were minor civil sub-divisions. While these statistical units are relatively comparable over time, they may be too large to be of value to a common small area system. There was little areal continuity between data collecting agencies; for example health units, polling districts, school divisions, manpower regions, travel and industry zones were not comparable to one another. In many instances, these data collection areas were too large extending beyond the entire AOSERP study region.

RECOMMENDATIONS

1. Of all the data inputs reviewed, the census provided the most comprehensive coverage of the target area. Therefore, it is recommended that Census data be given primary consideration in the initial development of the baseline data system.

2. Few of the statistical areas reviewed were comparable to one another or historically comparable over the last census decade. It is therefore recommended that, prerequisite to the establishment of a baseline data system, a common small area scheme be permanently delineated for the AOSERP target area.

3. It is further recommended that it would be premature to commence development of a baseline data system until the following conditions are satisfied:

   a. That a common small area scheme be adopted by local Provincial agencies and departments charged with the

responsibility of data collection.

b. That all demographic, social, economic, and biophysical data collected by local agencies and departments, as well as Statistics Canada, be geographically referenced to the common small area system.

It should be noted that while a baseline data system can be developed given the present status of data and statistical areas, its utilization would be hampered by the unevenness of data availability and lack of areal continuity between data collection agencies. The recommendations are made with the intent of satisfying the long-term goals of the project -- a traditional baseline data system with the capability of monitoring change in the AOSERP target area.

TABLE OF CONTENTS

Declaration ii
Letter of Transmittal iii
Descriptive Summary iv
STRUCTURE OF A TRADITIONAL BASELINE DATA SYSTEM ix
Conclusions and Recommendations xi

1. Review and Assessment of Data 1
1.1. Real Property Data System 1
1.2. Geographic Data System 2
1.3. Person Data System 4
1.4. Household or Family Data System 5
1.5. Common Small Area System 5
1.6. Recommendations 8

2. Data Evaluation 15
2.1. Demographic Data 16
2.1.1. Registered Live Births 17
2.1.2. Registered Marriages 18
2.1.3. Registered Deaths 18
2.1.4. Municipal Population Censuses 18
2.2. Social Data 19
2.3. Economic Data 20
2.4. Biophysical Data 21

3. Recommendations 22

4. Bibliography 24

LIST OF FIGURES

Figure 1. AOSERP study area location x
Figure 2. Enumeration areas, 1961 10
Figure 3. Enumeration areas, 1966 11
Figure 4. Enumeration areas, 1971 12
Figure 5. Minor civil sub-divisions, 1975 13
Figure 6. AOSERP study area and peripheral area boundaries 14

APPENDICES--SECTION 2

1.	Survey Response Evaluation	1
2.	Census Data File Listings, 1961, 1966, and 1971	45
3.	Demographic Data Contained in the Inventory of Statistical Files in the Alberta Government	59
4.	Social Data Contained in the Inventory of Statistical Files in the Alberta Government	69
5.	Economic Data Contained in the Inventory of Statistical Files in the Alberta Government	137
6.	Biophysical Data Contained in the Inventory of Statistical Files in the Alberta Government	165
7.	Data Availability for Fort McMurray	197
8.	Review of Existing Data Bank Systems	247
9.	Syncrude Canada Ltd. Reports	253
10.	List of Respondents and Questionnaire	261

NOTE: The Appendices have not been included with every copy of this report, because of their specialized interest. The body of the report outlines the conduct of the research and the conclusions drawn.

Any researcher wishing to review the appendices can peruse copies of the complete report available at the Alberta Environment Library, libraries of the University of Alberta, University of Calgary, and University of Lethbridge, and the Main Branches of public libraries in the cities of Alberta.

TRADITIONAL BASELINE DATA SYSTEM

In accordance with the research contract, dated 14 April 1976, this report completes the following proposal objectives:

1. Review and assessment of data.
2. Evaluation of weaknesses or deficiencies in existing data sources.
3. Recommendations based upon the foregoing evaluation.
4. Bibliography.

1. REVIEW AND ASSESSMENT OF DATA

As indicated in the proposal, a review and assessment of existing data sources pertinent to the AOSERP target area was conducted over the initial stage of the project. Several criteria were selected to assist in this evaluation of data. Fundamental to the development of a traditional baseline data system is to insure the existence of appropriate data inputs. To assist in the understanding of the selected criteria for evaluation purposes, however, it is necessary to briefly review the basic types of baseline data systems possible to develop.

A number of baseline data systems have been developed in North America and have been operational for a number of years. These data systems, while meeting a variety of needs in planning and policy making, are dependent upon the existence of advanced computer technology (hardware) and programs (software) designed to meet the user's needs through the manipulation and retrieval of appropriate data inputs. In order to accomplish this it is necessary that the data inputs be compatible with the baseline data system. Basically, there are four types of data systems that theoretically can be developed.

1.1 REAL PROPERTY DATA SYSTEM

In this system, the basic unit is the lot or land parcel. Data inputs or records allocated by lot would be such things as size of the lot, land use, zoning, land improvements, assessments, tax status, and so on. The identifying codes for the lot would be the address, lot number, block number; and in large municipalities,

census tract code, police beat, traffic zone, voting district, school zone, and planning district. This identifying information permits the use of data inputs collected by participating agencies, such as the local school administration, election office, transportation branch of city planning departments, and the various research sections of city planning. In summary, the specific content will vary with the use to which the system is put. With this system, the user can aggregate lot data to higher order areas such as the census tract or planning zone.

The Real Property Data System is the only type that has become operational in North America to any extent. This system meets the user needs of the large municipality where it is necessary to monitor on a weekly or monthly basis the changes that are occurring within the urban environment. Application of such a data system is enhanced by the fact that property assessment is a common function of cities, hence the only one satisfactory to meet a wide variety of uses. A local tax assessor's office maintains continuous records for each lot or land parcel within the municipality and usually is the most advanced in the mechanization of its records and procedures. It is also likely that other agencies charged with the responsibility of data collection to meet particular needs within the city will utilize the lot as their basic unit. Individual data linked to lot number or land parcel also alleviates difficulties encountered with the confidentiality issue. Most data used in such a system do not change with high rapidity which reduces problems of updating, i.e., on a daily or weekly basis.

This type of baseline system is operation in Detroit, Michigan; Washington, D.C.; New Haven, Connecticut; Alexandria, Virginia; and Philadelphia, Pennsylvania. Calgary, Alberta developed a similar data baseline system; however, the identifier is a roll number which is assigned to each parcel of land within the municipality. Other data content that can be referenced by roll number to the lot can therefore be integrated into the system.

1.2 GEOGRAPHIC DATA SYSTEM

In this system, the basic unit is similar to that of the

Real Property Data System, in that an areal unit is used. Normally, this areal unit is somewhat larger than the lot or parcel, e.g., a census geographical unit or a planning district. The record content correspondingly is different from that utilized in the aforementioned data system. In this system, the data content refers to the social events that occur on the land encompassed by the areal unit boundaries. Therefore, the data content must be geographically referenced, for example, crimes by residence or place of occurrence, welfare recipients by address, civil code violations by address, accidents by location, incidence of disease.

A Geographic Data System can provide ecological data for research, planning, program evaluation, and monitoring of changes within the area covered by the system. While the objectives of such a system meet a wide variety of needs, operationalization of a data system of this order is difficult. The reasons for this rest with the relatively more primitive state of record keeping or mechanization of original records within various agencies; the difficulty in relating address coded records to the basic areal unit used in the data system; and the problems associated with classified or confidential information. Utilization of data records already aggregated to usable areal units, i.e., areal units compatible to the common small areas used in the data system, eliminates the confidentiality problem associated with individual data.

Washington, D.C. has generated a Geographic Data System for monitoring changes in an urban reneval target area with the overall purpose of detecting early symptoms of deterioration. The smaller target area permitted the planners to accommodate address-coded data more readily; however, address-coded data would be a tedious and time-consuming task for larger target areas.

Another example of the Geographic Data System is the Population Research Laboratory's SUDS system (Standard User Data Sets) which uses the census enumeration area as the basic unit. Acquisition of the census magnetic tapes for 1961 and 1966 permitted the development of retrieval and manipulation programs for accessing the tape files. The objective of the SUDS system was to generate census

data profiles for any user-designated area in Canada. The requirements were that the user's designated area be any combination of census enumeration areas and that the data content conform to the standard data sets provided by the system. All of the 1961 and 1966 census data were identified by enumeration area. With the enumeration area as the basic building block, the data system can retrieve and manipulate data to higher order aggregations, such as census tract, metropolitan area, county, municipality, census division, and other minor sub-divisions recognized by the census. Since most data requests were for target areas within Alberta, an "Alberta Only" tape file was constructed which, in turn, reduced computer execution time in meeting data needs. Formerly, requests for data concerning parts of Alberta necessitated scanning all of the tape files (27 magnetic tapes), whereas the Alberta Only files consisted of three magnetic tapes.

1.3 PERSON DATA SYSTEM

In this system, the basic unit is the individual. The content of the record consists of selected characteristics or behavior sets about the person as known to participating agencies. Examples of such agencies could be the Hospital Commission, Provincial Welfare Department, Motor Vehicles, Juvenile Court, etc. The data system then consists of all persons for whom a record can be established. The code identifier would be a number, such as a social insurance number, file number, or tax number. All personal information would then be aggregated according to the individual's identification number. In essence, a personal history or dossier is maintained for computer access. This type of data system would have to be address-linked in order for it to be relevant to a particular geographic area. An additional difficulty encountered with this approach is that of confidentiality of individual record files from various provincial agencies. Generally, name or address-linked records are considered to be classified information, and therefore its availability would be severely restricted. The best example of this form of data system would be credit rating data banks where historical consumer information, payment records, and personal behavior patterns are stored on an individual basis. It would be assumed

that the Internal Revenue Service in the United States, as well as the Receiver General for Canada would have access to a Person Data System in terms of historical tax files.

1.4 HOUSEHOLD OR FAMILY DATA SYSTEM

The basic unit is the household or family. This data system is similar to the Person Data System, with the exception that the data are linked to the larger family or household unit rather than the individual. The difficulties encountered in operationalizing such a system, however, would be the same. For example, problems of confidentiality and address coding would apply.

It is important to recognize the objectives set forth in the proposal which entail the establishment of a basic data system with the capability of monitoring upon demand change for any user-designated area within the AOSERP target area. These objectives can best be met by the Geographic Data System described above. All data inputs incorporated in such a system would have to be linked to the basic geographical units used. This type of data system is further justified, in that the problem of restricted or confidential data is eliminated. In other words, the confidentiality of individual records is not violated since data aggregations only will be used. This system also reduces the necessity to link or match addresses to the basic geographic unit. Data profiles, whether demographic, social, economic, or biophysical, can be retrieved from such a system for user-designated areas with relative efficiency. In addition, the system would be amenable to standardized statistical techniques available through SPSS, as well as computer mapping programs such as SYMAP and CALCOMP. These graphics computer packages would enable the user to retrieve graphic profiles for selected areas within the AOSERP region, thus providing the user with an additional analytical tool for data interpretation.

1.5 COMMON SMALL AREA SYSTEM

Fundamental to this type of baseline data system is the existence of a usable small areal unit scheme that is common to a variety of data sources. For the purpose of monitoring change in the AOSERP region, it is necessary that the basic areal unit be small

enough to afford a reasonably fine grain for analytical purposes. In other words, the jigsaw puzzle pieces making up the target area shouldn't be too large.

In the establishment of a baseline data system utilizing geographic statistical areas as the basic building blocks, a fundamental problem becomes evident. Areal continuity becomes the prerequisite, in that data inputs from various sources would have to be linked to the common small area system.

In the evaluation of the variety of data sources available for the AOSERP region it was found that the basic deficiency was lack of areal continuity between the statistical areas established by various governmental departments and agencies. It is recognized that each agency or department develops their own statistical area system designed to meet specific needs. Not only is there inter-departmental variation in statistical area schemes, but also temporal variation, in that some departments modify their areal schemes over time as their needs or objectives change.

It was found that statistical areas vary substantially in size depending upon the nature of the data required by the investigative group or agency. For the AOSERP region, statistical areas by which data are available can vary in size from a sub-division of Fort McMurray to an entire census division or health unit encompassing the entire target area. Between these two extremes there are numerous sizes of statistical areas including Indian reserves, enumeration areas, polling districts, counties, municipal districts, and municipalities.

To demonstrate this lack of areal continuity, it was found that the most common statistical area compatible to the AOSERP region (Census Division 12) could not be matched by the areal units delineated by other provincial departments. In this example, the statistical areas of nine government agencies are compared to the spatial extent of Census Division 12:

-- Department of Advanced Education and Manpower - Canada Manpower Areas, 1974: In 1974, Census Division 12 comprised the Fort McMurray and St. Paul Canada Manpower Areas.

-- Department of Agriculture - Agricultural Regions, 1974: In 1974, the Northeast Agricultural Region included Census Division 12 and parts of Census Divisions 10 and 7.

-- Department of Agriculture - Agricultural Reporting Areas, 1975: In 1975, ARA 6 included not only Census Division 12, but also Census Divisions 13 and 14.

-- Department of Municipal Affairs - Minor Civil Sub-Divisions, 1975: In 1975, Census Division 12 was compatible with Improvement Districts 18 and 24; Counties 13 and 19; and Municipal District 87.

-- Department of Social Services and Community Health - Health Units, 1976: In 1976, Census Division 12 contained the Fort McMurray and District Health Unit, the Northeastern Alberta Health Unit, and part of Athabasca Health Unit.

-- Electoral Districts, 1971: In 1971, Census Division 12 comprised parts of Electoral Districts 801, 818, and 815.

-- Industry and Commerce Regions, 1976: In 1976, Industry and Commerce Region 6 includes that portion of Census Division 12 south of the 19th base line and part of Census Division 13.

-- Judicial Districts, 1962: In 1962, the Edmonton Judicial District included all of Census Divisions 12, 13, 14 and parts of Census Divisions 10, 11, and 15.

-- Travel Alberta - Travel Industry Zones, 1973: In 1973, 'The Lakeland Travel Zone' included all of Census Division 12 and parts of Census Divisions 10 and 13.

As can be seen from the foregoing example, there is limited comparability between the various statistical area schemes established by the nine governmental agencies and Census Division 12. It should be noted that while the Canada Manpower Areas and the Minor Civil Sub-Divisions were compatible with Census Division 12, the areal scale of these statistical units was large enough to cast doubt upon their utilization in a common small area system.

The only data collection agency that used a relatively small areal unit scheme was Statistics Canada. This governmental agency utilizes the enumeration area as the basic data collection unit. For the purpose of evaluation, it was determined that Census Division 12 approximates the AOSERP region. This census division encompasses AOSERP study area and most of the peripheral area defined by the Human

Environment Committee boundary. Three maps were prepared showing the 1961, 1966, and 1971 enumeration areas that are contained within this census division. (Figures 2, 3, and 4)

The enumeration area as used by the census is essentially an administrative unit; however, all census data are linked to this statistical area. Enumeration area data are available on census summary magnetic tapes and selected variables are provided by enumeration area in special tabulations. The major limitation associated with the enumeration area is that its boundaries are changeable, hence there is no guarantee of areal comparability from census to census. This limitation can be seen in the following charts. The optimum population size of the enumeration area ranges between 250 and 300 households. As the population changes, the enumeration area changes accordingly to maintain the optimum size. Regardless of this limitation, the enumeration area system affords the finest grain in areal scale available for the target area.

Figure 5 portrays the 1971 enumeration areas for the Town of Fort McMurray. To illustrate the changeability of enumeration areas, there were two enumeration areas in this town for 1961, and in 1966, there were four enumeration areas. The number of enumeration areas had increased to eight by 1971 reflecting the dynamic growth of Fort McMurray over the 1961 to 1971 period. Similar changes in enumeration areas occurred in the southern part of the census division. This lack of areal continuity over time in enumeration areas severely limits the monitoring of historical changes in the available census data sets.

1.6 RECOMMENDATIONS

Considering the difficulties with lack of areal continuity between departmental statistical schemes and the temporal variations found in enumeration areas, two recommendations are advanced to resolve these limitations.

1. <u>It is recommended that Statistics Canada be requested to standardize the enumeration areas in Census Division 12 according to 1976 delineations</u>. This would insure areal continuity for future censuses; however, the problem of areal continuity in past censuses

would remain.

2. Given the above recommendation, it is further suggested that the Province encourage the various departments and agencies in the government to utilize these standardized enumeration areas as the basic building blocks for their own statistical areas.

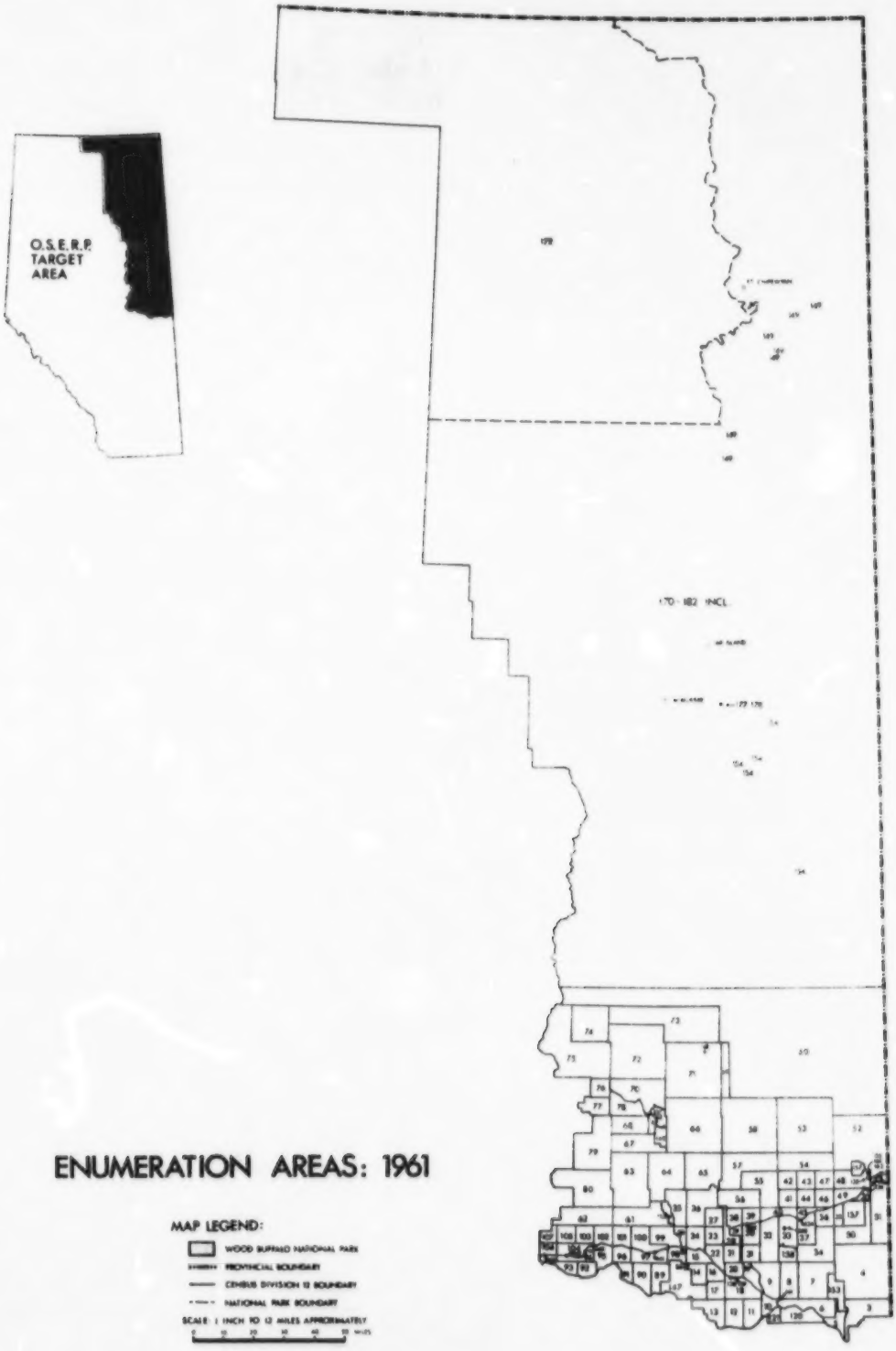

FIGURE 2. Enumeration areas, 1961.

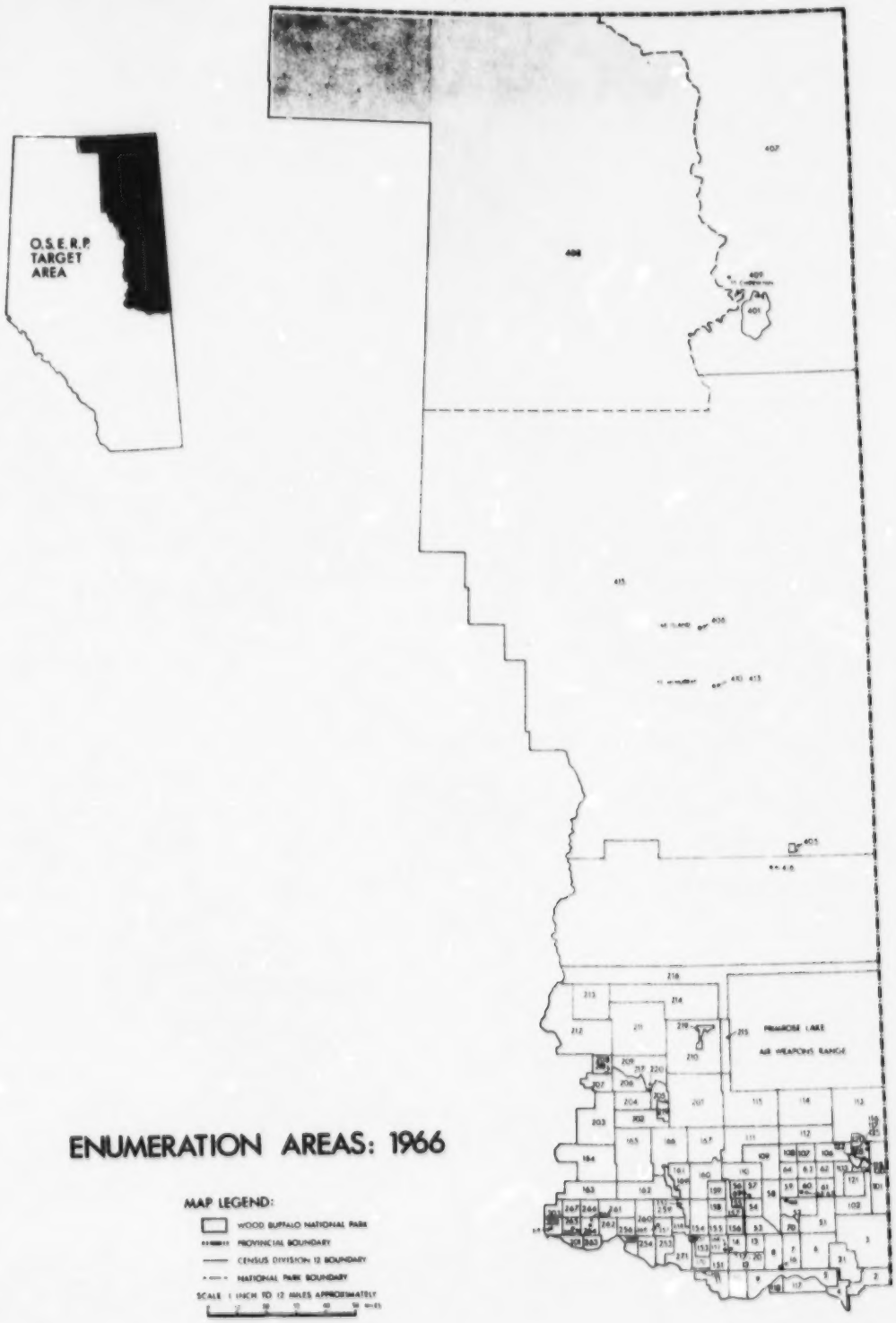

FIGURE 3. Enumeration areas, 1966.

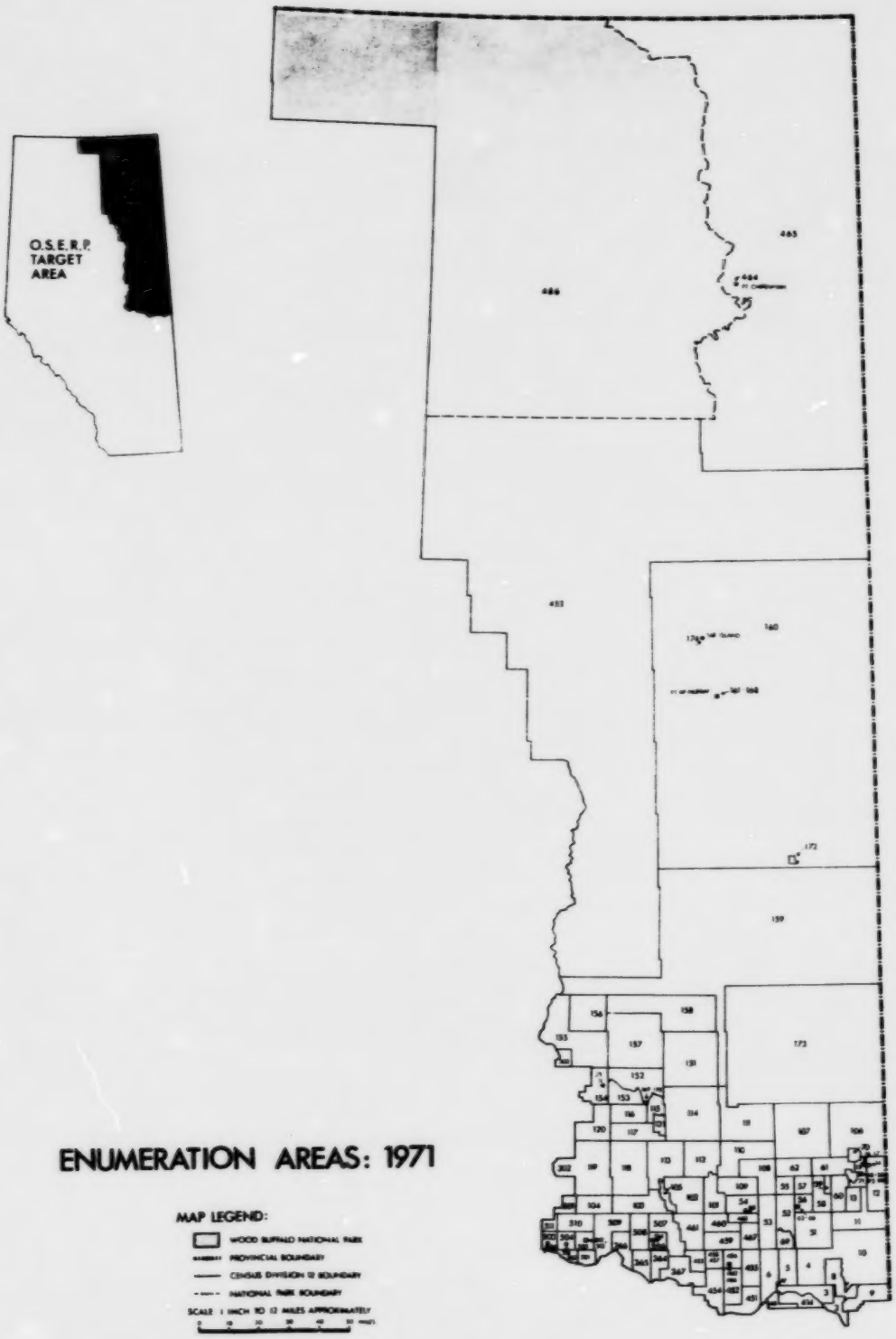

FIGURE 4. Enumeration areas, 1971.

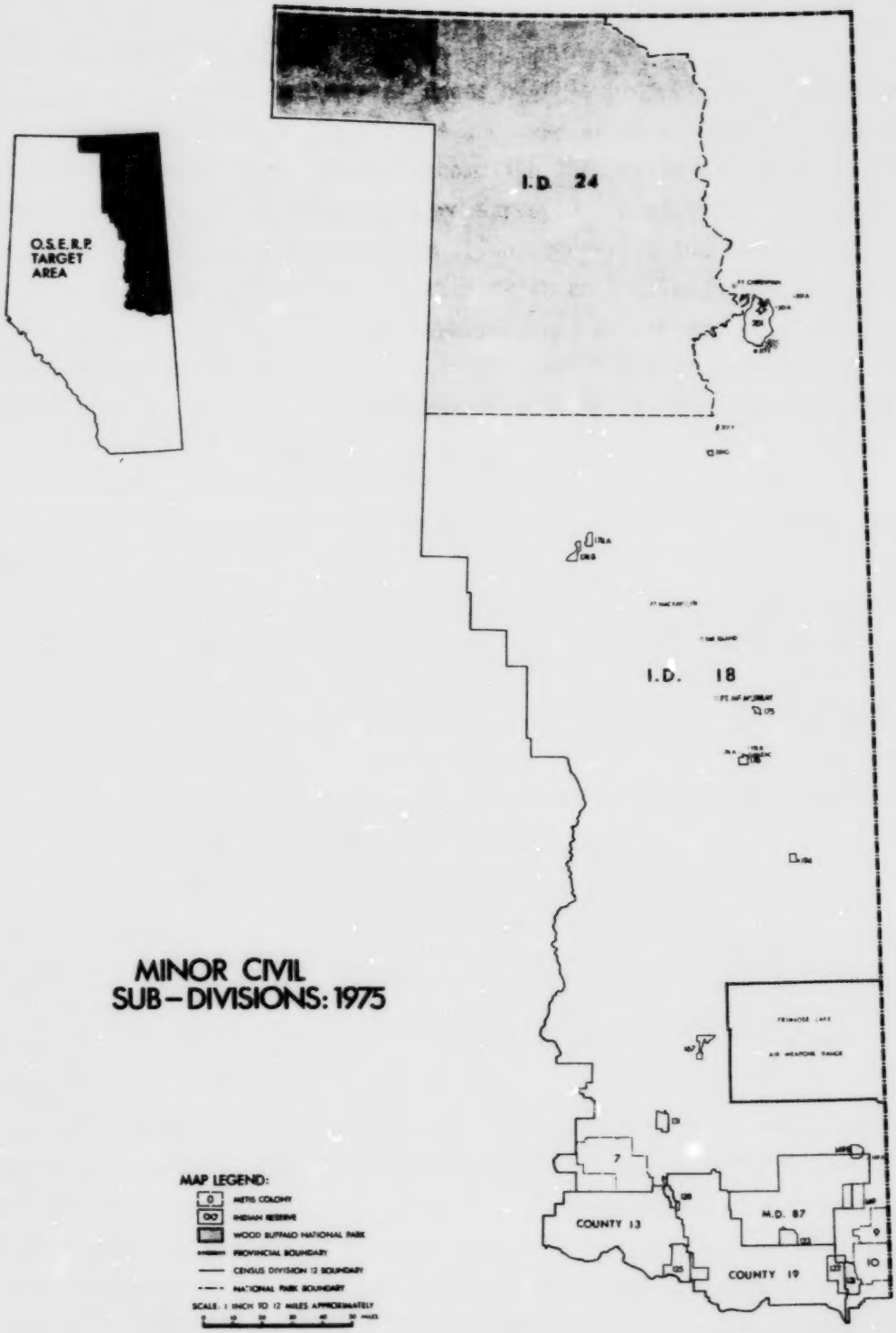

FIGURE 5. Minor civil sub-divisions, 1975.

FIGURE 6. AOSERP study area and peripheral area boundaries.

2. DATA EVALUATION

The second fundamental phase of this project was to evaluate the various types of data available and make an assessment in terms of data deficiencies. In order to accomplish this task, all agencies and departments within the Province were surveyed. A questionnaire was designed requesting specific information concerning the type of data collected, its accessibility, continuity in collection, and geographic reference. Fifty-nine agencies were surveyed and thirty-nine responses have been received and evaluated providing a 66 per cent response rate. A listing of respondents and sample of the questionnaire are included in the appendix. In several instances, statistical compilations were provided along with the completed questionnaire. Samples of these compilations appropriate to the AOSERP region are also included.

The criteria selected for evaluation of data are as follows:

1. Type of Data Collected. Demographic, Social, Economic, or Biophysical
2. Data Source. Name of agency responsible for data collection.
3. Data Detail. Summary of types of primary data collected, how tabulated, and extent of detail.
4. Geographic Identification. Areal reference of data, i.e., statistical unit.
5. Accessibility. Availability and confidentiality status of data.
6. Data Time Frame. Initiation of data collection and regularity of collection.

These criteria permitted the assessment of data compatibility to a Geographic Data System, in that it is necessary to determine whether data inputs would be amenable to geographically coded continuous monitoring systems. The data must be accessible without violation of confidentiality; must be available in aggregate form; must be regularly collected to insure continuity; and must be conceptually consistent.

The assessment of data will appear in order of classification,

i.e., demographic, social, economic, and biophysical data. Following these general assessments, there will be more specific data evaluation forms with comments according to the aforementioned criteria.

2.1 DEMOGRAPHIC DATA

By far, the greatest bulk of demographic data for the AOSERP target area is provided by Statistics Canada through the data collection vehicles of the census and vital registration. The Canadian government's large-scale census operation is intended exclusively to generate information. There is no counterpart of these massive operations existing at the local and provincial level. Provincial department data collecting efforts are particularistic, in that they are by-products of operating programs. The direction of local data collection efforts will change according to shifts in local objectives, budgets, expertise, and policy.

Census data are collected on a decennial basis in every year ending in one, e.g., 1951, 1961, and 1971. A more limited census is conducted every five years in years ending in six, e.g., 1956, 1966, and 1976. The smallest areal unit by which census data are available is the enumeration area. The census enumeration area data are in unpublished forms, such as special tabulations for selected census variables and Census Summary Magnetic Tapes. Tape file descriptions appear in the appendix revealing the comprehensive detail of demographic, social, and economic data available for 1961, 1966, and 1971.

Census data are available in simple tabulations, as well as cross-tabulations. Cross-tabulations are distinguished from single tabulations, in that they show more than one variable, e.g., age and level of education. These data are accessible in a form compatible to a Geographic Data System. They are collected on a regular basis and are geographically coded. The concepts and definitions utilized in the census are generally consistent over time. In addition, the areal coverage is complete for each census year.

The limitations of census data are several. First, while the collection of data is conducted on a regular basis, the five year time span between censuses is too long particularly concerning areas that experience dynamic population change. Secondly, there is

a time lag encountered between the collection of census data and the time when the data becomes available to the user. For monitoring short-term changes in a target area, it would be necessary to find reliable local sources to supplement and up-date census data. In light of the previous assessment of statistical areas utilized by local agencies -- lack of areal continuity -- resolving this problem would be extremely difficult. While census data are most adaptable to a baseline data system, a serious sacrifice in frequency would have to be made.

The Vital Statistics Division of the Provincial Department of Social Services and Community Health serves as a collection agent of vital statistics at the local level. The vital registration process entails the collection of data concerning change in status at the individual level. These vital data are then aggregated and provided by municipality, minor civil sub-divisions, census divisions and province. The vital statistics available for the above geographic areas are live-births by sex, legitimate and illegitimate births, stillbirths, deaths, infant deaths, and maternal deaths. In addition, deaths by cause and sex are available in tabulated form at the census division level only. The definitions and concepts utilized have been standardized so that they are consistent over time.

The limitation of vital statistics is the time lag between collection date and release of data. Since the vital registration is an ongoing process, vital statistics reports are released on an annual basis, however, there is approximately a two-year lag. The data content is valuable, in that measurement of natural increase, morbidity, and mortality is attainable.

Various departments and agencies of the Provincial government also collect and process data on an annual basis. These data would be of use in supplementing the census data over the intercensal periods. The types of information and responsible departments and agencies collecting data on an annual basis are as follows:

2.1.1 Registered Live Births

Number of registered live and stillbirths. Livebirths are provided by month of the year, age of the mother, birth weight of the

infant, etc. Stillbirths are provided for the period of gestation, by sex of infant, and by age of mother. This information is in tabular form by place of birth and residence of mother. (Alberta Social Services and Community Health - Division of Vital Statistics)

2.1.2 Registered Marriages

Information includes the place and date of marriage, as well as the marriage license number. In addition, the name, social insurance number, previous marital status, religion, birthdate, and place of residence are provided for both bride and groom. (Alberta Social Services and Community Health - Division of Vital Statistics)

2.1.3 Registered Deaths

Deaths by cause, sex, marital status, and month are tabulated for census divisions and urban centers. (Alberta Social Services and Community Health - Division of Vital Statistics)

2.1.4 Municipal Population Censuses

Population by type of municipality, e.g., village, town, city, minor civil sub-divisions, is listed. (Alberta Municipal Affairs - Municipal Inspection and Advisory Services Branch) It is likely that actual civic census figures are available only for Fort McMurray. Population estimates based on the previous federal census undertaking are used for the remaining communities in the AOSERP target area.

The Provincial government Departments of Social Services and Community Health and the Alberta Health Care Insurance Commission also collect and process general demographic data. Alberta Social Services and Community Health have such vital statistics as livebirths, deaths, morbidity data, marriages, and age-sex distributions matched by postal code. These data can be aggregated to the census division level.

In summary, federal census statistics are often inadequate for a monitoring system since they have considerable time lags between collection and availability. The time lag difficulty is somewhat reduced when using vital statistics. The census data are available for small geographic areal units, however, these areal units are not standardized over time. Vital registration data are more frequently available, but are geographically referenced to larger statistical

units. The wealth of consistent demographic data for the AOSERP region may offset the geographical timeliness limitations. These limitations notwithstanding, the census and vital registration are the best available sources of geographic-referenced population data. Data collected by local and provincial agencies have limited use, in that they are irregularly collected, have limited accessibility, inconsistent definitions and conceptualization, and are byproducts of specific organizational goals.

2.2 SOCIAL DATA

It should be noted at this point that the census and vital statistics sources provide data classified as social and economic, hence, the preceding statements have relevance in this section. In addition, a diverse range of government departments and agencies collect and process information which may be referred to as social statistics. A list of the social data maintained by various departments and agencies taken from the Inventory of Statistical Files in the Alberta Government appears on the following pages. Detailed descriptions of each data set appear in the appendix.

Most of these data are generated for intradepartmental purposes and, therefore, are of limited value to the AOSERP project. At best this diverse range of data provide some measurement of questionable validity and reliability.

Of the information recorded in the Inventory, the following data collected by the Alberta Housing Corporation would be of special interest: residential land inventory, the apartment accommodation survey, and the market rent survey of single-detached dwellings.

In the departmental responses to the questionnaire, the Alberta Department of Social Services and Community Health processes the following data:

a. notifiable diseases.
b. reported poisonings.
c. handicapped registry, including mental and physical disabilities.
d. public assistance recipients.
e. child welfare cases.
f. mental health patients.

The geographic references for these social and health data are the Alberta Health Units and minor civil sub-divisions. For data listed in the Inventory and the individual evaluations, accessibility is usually restricted since they are personal record files. Utilization of individual records for a data system requires tedious manipulation of address information so that the data can be aggregated to a higher order geographical unit for analysis. The main limitation of the social data, as indicated in the appendices, is that of confidentiality.

2.3 ECONOMIC DATA

In addition to the economic statistics collected by Statistics Canada, various governmental departments collect economic data for in-house use. A list of these various economic statistics appears in the appendix under the heading Inventory of Statistical Files in the Alberta Government - Economic Data. A complete explanation of each data set appears in this appendix. Of this listing of data, the summaries of property assessments and the municipal financial and statistical statements would be of most value to the Geographic Data System.

The municipal financial and statistical statements are made available through Alberta Municipal Affairs - Municipal Inspection and Advisory Services Branch and include the following categories of data:

a. financial statements.
b. analysis of debenture debts.
c. property assessments.
d. valuation of property exempt from taxation.
e. transfers receivable.
f. grants in lieu of taxes receivable.
g. general population and geographic statistics.

The above listed data are aggregated for villages, towns, cities, minor civil sub-divisions, and census divisions. The data are collected on an annual basis with approximately a two year time lag between collection and release.

The summaries of property assessments processed by the Alberta Municipal Affairs - Assessment Equalization Board are classified as restricted information, however, access may be achieved through

appropriate channels. The data collected and processed by this agency include assessment values on privately owned property, municipally owned property, property owned by the province, property owned by the Federal government, and tax exempt property.

2.4 BIOPHYSICAL DATA

Alberta Energy and Natural Resources - Renewable Resources (Lands and Forests) and Alberta Environment collect the largest proportion of the biophysical data available in the Province. A list of the biophysical data appears in the appendix. These data are largely based on irregular surveys conducted by the two departments. The areal extent and sample size between studies varies substantially as does the frequency of up-dating the information.

The Land Use Assignment and Renewable Resources Sections of Alberta Energy and Natural Resources also provide some biophysical data. The Land Use Assignment Section has been responsible for the publication of two reports dealing with renewable natural resources and the capability of land to produce them. The geographical extent of these studies is limited to the Fort McMurray - Gregoire Lake Areas and the Namur Lake Area. The Renewable Resources Section has produced map sheets pertaining to the AOSERP region at a scale of 1" to 1 mile (1:63,360) on which the following data have been plotted: density, height and species of tree stands, detailed access, oil and gas fields, transmission lines, geophysical activity, ranger stations, air strips, recreation areas, parks, Indian reserves, Metis colonies, and detailed hydrology.

Additional biophysical information has also been published in a series of reports by Syncrude of Canada Limited. The types of information in these reports is noted in the annotated bibliography, however, it should be noted that the areal extent of these reports is limited to the immediate Syncrude Lease Areas, i.e., they are site-specific.

The biophysical data available are geographically referenced and can be manipulated to match the basic units of the data system. It is noted, however, that several of the surveys are site-specific and have been conducted on an irregular basis.

3. RECOMMENDATIONS

In consideration of the study objectives, the most reliable and regularly collected data amenable to the baseline data system are provided by Statistics Canada. The census is the most comprehensive source covering the demographic, social, and economic classifications. Its chief limitation is the time frame for data collection and time-lag in data release. The inter-censal deficiencies could be supplemented by locally collected data, however, there has been little consistency and scope in the data collection at the local level. Locally collected data are of limited value in terms of geographic reference to the AOSERP target area, i.e., the local data are collected for much larger areas than the target area, or the areas used are not small enough to be of value. This is the severest limitation of the social and economic data. There is, in addition, extreme difficulty in assessing the quality of data collected as well as consistency in conceptualization. In many instances, the respondents did not address themselves to this question in the questionnaire or, because of poor documentation in data collection, simply did not know how variables were defined or if definitions had changed. Utilization of selected data from such sources would entail careful scrutiny through face validity and content validity checks prior to integration into the system.

The data area of interest which requires the greatest improvement is in the availability of biophysical data. The data available through various agencies and departments of the provincial government are usually time-specific (i.e., one shot surveys) and are collected on the basis of in-house objectives. There are also substantial gaps in the types of information made available. This situation is hardly conducive to the establishment of a baseline data system with the intent of monitoring environmental changes.

The general recommendations concerning data availability emerging out of this study are as follows:

1. In considering the weaknesses and advantages of the various demographic, social, and economic data inputs, it is recommended that the Census and Vital Registration sources be used for the basic data inputs.
2. It is further recommended that local data collection agencies

be used to supplement the above sources in the intervening years. In addition, local data should be upgraded and properly documented before selection as inputs for the system.

3. It is further recommended that the types of information available for the town of Fort McMurray be collected for all settlements in the AOSERP region. (see appendix)

4. BIBLIOGRAPHY

Baird, John C. 1972. Application of Information Theory to Man Environment Systems. J. Amer. Inst. Planners. 38 (4): 254-257.

Baum, Martin A., David F. Bergwall, and Phillip N. Reeves. 1975. Planning Health Care Delivery Systems. Amer. J. Public Health. 65 (3): 272-279.

Bilsborrow, Richard E. 1974. The Preparation of an Inventory of Demographic Data for Social and Economic Planning. Laboratories for Population Statistics, Manual Series No. 5, July 1974.

Brounstein, Sidney H. 1968. Some Concepts and Techniques for Constructing and Using a Geographically-Oriented Data Base. Socio-Economic Planning Sciences. 1(November): 309-325.

Churchman, C. West. 1968. The Systems Approach. New York: Delta.

Clark, I.N. 1973. Community Social Indicators. Urban Affairs Quart. 9(1): 3-37.

Conn, W. David. 1975. The Difficulty of Forecasting Ambient Air Quality: A Weak Link in Pollution Control, J. Amer. Inst. Planners. 41(5): 334-346

Friedmann, John. 1968. An Information Model of Urbanization, Urban Affairs Quart. 4.

Hansen, Niles M. 1965. Regional Planning in a Mixed Economy, Southern Econ. J. October 1965.

Harris, Britton. 1966. The Uses of Theory in the Simulation of Urban Phenomena, J. Amer. Inst. Planners. 32 (5).

Hearle, Edward F.R. 1968. Urban Management Information Systems, Socio-Economic Planning Sciences. 1 (November): 215-221

Hirschman, H.O. 1967. Development Projects Observed, Chapter 1. Brookings Institution, Washington, D.C.

Ide, Edward, A. 1970. Address Coding to Produce Age-Race Data by City Block for School Planning, Socio-Economic Planning. 4(March): 107-118.

Krotki, Karol. 1969. Availability and Retrieval of Regional Data in Canada. In B.Y.Card (Ed.) Perspectives on Regions and Regionalism Western Ass. Sociology and Anthropology Proc. Banff, Alberta, December 28-30, 1968. Edmonton, Alberta.

Lyle, John and Mark von Wodtke. 1974. Information System for Environmental Planning, J. Amer. Inst. Planners, 40 (6): 394-413.

Lyon, C.S. 1969. Guide for Users of Enumeration Areas Print-Out Data From the Census of Population and Housing 1961 and 1966. Dep. Sociology, Univ. Alberta, Edmonton, Alberta.

Mayer, Robert R. 1972. Social System Models for Planners, J. Amer. Inst. Planners. 38(3): 130-139

McKean, Roland N. 1963. Efficiency in Government Through Systems Analysis. John Wiley & Sons, New York.

McVey, W.W. 1970. Development of a Profile Facility for User Designated Areas. Population Res. Lab., Dep. Sociology, Univ. Alberta, 8 August, 1970.

_____. 1971. Standard User Data Sets. Population Res. Lab., Dep. Sociology, Univ. Alberta, Pop. Memo 25, 8 June 1971.

Muhlin, G.L. and B.I. Milcarek. 1974. Urban Analysis and Planning: A Cautionary Note on the Utilization of Census Data, Urban Affairs Quart. 10(2): 212-222.

Nelson, Claron E. 1967. Resource Information for Area Research, The Annals of Regional Science. 1 (1): 7-22.

Ramey, James W. 1968. Urban Data Banks and the Rights of the Individual, Socio-Economic Planning Sciences. 1 (November): 327-333.

Savas, Emanuel S. 1968. Information Systems in a New York Urban Observatory, Socio-Economic Planning Sciences. 1 (November): 203-208.

Schon, D.A. 1968. Forecasting and Technological Forecasting. In Toward the Year 2000. Daedalus.

Shaw, Paul R. 1971. A Record Linkage System as a Means to Extensive Socio-Economic Farm Population Data: 1971 Census of Canada, Socio-Economic Planning Sciences 5 (June): 221-230.

Snider, E.L. 1972. Toward the Development of a Socio-Political Data Bank for Alberta, Alberta Human Resources Research Council, Edmonton, Alberta.

Steger, Wilbur R. and T.R. Rakshmanan. 1968. Plan Evaluation Methodologies - Some Aspects of Decision Requirements and Analytic Response. In George C. Hemmens (Ed.) Urban Development Models. Highway Research Board, Washington, D.C.

Teitz, Michael B. 1968. Cost-Effectiveness: A Systems Approach to the Analysis of Urban Services, J. Amer. Inst. Planners, 34 (Sept.).

Tropman, John E. 1969. Critical Dimensions of Community Structure. Urban Affairs Quart. 5 (2): 215-232.

U.S. Bureau of the Census. 1966. GE 40, No. 1, 1965, Papers Presented at the Census Tract Conference, September 10, 1965. Washington, D.C.

U.S. Bureau of the Census. 1970. GE 40, No. 6, Final 1970 Census Plans and Four Programming Systems for Computerized Data Retrieval and Manipulation, New York, N.Y., August 21, 1969. U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census. 1968. GE 40, No. 4, Papers Presented at the Conference on Small Area Statistics, Amer. Stat. Ass., Washington, D.C. December 1967. U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census. 1968. GE 40, No. 5, Papers Presented at the Conference on Small Area Statistics, Amer. Stat. Ass., Pittsburg, P.A., August 23, 1968. U.S. Government Printing Office, Washington, D.C.

U.S. Bureau of the Census. 1975. A Numerator and Denominator for Measuring Change. U.S. Department of Commerce, U.S. Government Printing Office, Washington, D.C.

Veeder, Nancy W. 1975. Health Services Utilization Models for Human Services Planning, J. Amer. Inst. Planners. 41 (2): 101-109.

von Bonventer, E. 1964. Spatial Organization Theory as a Basis for Regional Planning, J. Amer. Inst. Planners. 30 (May).

Wildavsky, Aaron. 1966. The Political Economy of Efficiency: Cost Benefit Analysis, Systems Analysis, and Program Budgeting, Public Admin. Rev. 26 (Dec.).